

Perturbation Analysis and Randomized Algorithms for Large-Scale Total Least Squares Problems

Pengpeng Xie^{a *} Yimin Wei^{a,b†} Hua Xiang^{c‡}

^a School of Mathematical Sciences, Fudan University, Shanghai, 200433, P.R. China

^b and Key Laboratory of Mathematics for Nonlinear Sciences

^c School of Mathematics and Statistics, Wuhan University, Wuhan, 430072, P.R. China

November 12, 2014

Abstract

In this paper, we present perturbation analysis and randomized algorithms for the total least squares (Tls) problems. We derive the perturbation bound and check its sharpness by numerical experiments. Motivated by the recently popular probabilistic algorithms for low-rank approximations, we develop randomized algorithms for the Tls and the truncated total least squares (Ttls) solutions of large-scale discrete ill-posed problems, which can greatly reduce the computational time and still keep good accuracy.

Keywords: Condition number; Singular value decomposition; Total least squares; Truncated total least squares; Randomized algorithms.

AMS Classification: 15A09, 65F35

*E-mail: 09110180005@fudan.edu.cn. P. Xie is supported by the National Natural Science Foundation of China under grant 11271084.

†E-mail: ymwei@fudan.edu.cn. Y. Wei is supported by the National Natural Science Foundation of China under grant 11271084.

‡Corresponding author (H. Xiang). E-mail: hxiang@whu.edu.cn. H. Xiang is supported by the National Natural Science Foundation of China under grants 10901125 and 91130022. Partial work was completed when he visited Fudan University and The Chinese University of Hong Kong in 2014.

1 Introduction

Given an overdetermined set of m linear equations $Ax \approx b$ in n unknowns x , the total least squares (TLs) problem can be formulated as [38]

$$\min \| [E \ f] \|_F \quad \text{subject to} \quad b + f \in \mathcal{R}(A + E), \quad (1.1)$$

where $\|\cdot\|_F$ denotes the Frobenius matrix norm and $\mathcal{R}(\cdot)$ represents the range space. When the sampling or modeling or measurement errors also affect the coefficient matrix A , the TLs method is more realistic, while the underlying assumption in the least squares (Ls) problem is that errors only occur in the right-hand side vector b .

The term “total least squares” was coined in [11]. It has been also known as errors-in-variables model, orthogonal regression, or measurement errors in the statistical literature, and blind deconvolution in image deblurring. In the monograph [38] the authors show the readers how to use TLs for solving a variety of problems, especially those arising in signal processing, medical imaging, and geophysics, etc. The applications and theory associated with the TLs are still being studied, for example [16, 23]. In recent years, perturbation analysis for the TLs problem has been studied extensively in the numerical linear algebra (see e.g. [2, 6, 7, 13, 17, 24, 28, 31, 33, 40, 39, 43, 44]).

It is well known that the condition number indicates the sensitivity of the problem itself, and that an approximate bound for the forward error can be given by the multiplication of the condition number and the backward error. For the perturbation in the solution of the scaled total least squares problem, Zhou et al. [44] presented a first order estimation. But as pointed out by the authors, it is not easy to compute since the condition number formula is a Kronecker product-based one. Baboulin and Gratton [2] derived a computable expression for the condition number. At almost the same time, Li and Jia [24] made a first order perturbation analysis. Recently, Jia and Li [18] proposed a formula which only used the singular values and the right singular vectors of $[A, \ b]$, and presented the lower and upper bounds for the condition number. In this paper, we will present a relative perturbation bound without considering the condition number only. We first give a perturbation bound in this paper. And its significant improvements will be demonstrated by numerical examples. We also show that these three condition numbers in [2, 24, 44] mentioned above are mathematically equivalent.

For the numerical solution of the TLs problem, a simple and elegant solver based on the SVD of the augmented matrix $[A, \ b]$ can be used. When A is large, a complete SVD will be very costly. One improvement is to compute a partial SVD based on Lanczos bi-diagonalization [9]. But the partial SVD is still prohibitive for large-scale sparse or structured matrices, since the initial reduction of $[A, \ b]$ to bi-diagonal form will destroy the sparsity or structure of the matrix. For the TLs problem with very ill-conditioned coefficient matrix whose singular values decay gradually, the task is even more challenging. Without regularization, the ordinary least squares or total least squares solvers yield physically meaningless solutions. For such discrete ill-posed problems, there already exist several regularization strategies of the TLs solution. For example, the solution can be stabilized by truncating small singular values of $[A, \ b]$ via an iterative algorithm based on Lanczos bi-diagonalization [9]. Tikhonov regularization strategy is used in [4, 10, 22, 23, 29], where a Cholesky decomposition is computed in each step in [4], and the linear systems are projected onto Krylov subspace of much smaller dimensions to reduce the problem size in [22]. Regularization by an additional quadratic constraint is another choice [5, 21, 35, 36], which is the regularized TLs based on quadratic eigenvalue problems (QEP): adding

a quadratic constraint to the TLS, and then iteratively solving the QEP. For the large-scale discrete ill-conditioned problem, a complete SVD is prohibitive, and the choice of regularization parameter is also time consuming. The classical SVD of a matrix can be well approximated by the randomized SVD [14], and the regularization parameter can also be located by randomized algorithms [42]. Such randomized algorithms can greatly reduce the computational time, and still keep good accuracy with very high probability. Motivated by these randomized matrix algorithms, we present randomized algorithms for the solution of total least squares (TLS) problems including the well-conditioned cases and the ill-conditioned cases. For the practical cases where the numerical rank is not known, randomized algorithms can usually be implemented in an adaptive approach with the sample number increasing until the desired tolerance is satisfied. Here the tolerance parameter is adopted to describe how well the basis matrix generated by the randomization captures the action of the target matrix. Based on this, we develop the randomized algorithm for the fixed precision cases.

Throughout this paper, $\mathbb{R}^{m \times n}$ denotes the set of $m \times n$ matrices with real entries and I_n stands for the identity matrix with order n . As usual, $\mathbf{0}$ denotes the zero matrix with the corresponding size easily known from the context. For a matrix $A \in \mathbb{R}^{m \times n}$, A^T is the transpose of A ; $\|A\|_2$, $\|A\|_F$ and $\|A\|_\infty$ denote the spectral norm, the Frobenius norm and the infinity norm of A , respectively. And A^\dagger represents the Moore-Penrose inverse of A [12] and $\lambda_{\max}(A)$ denotes the largest eigenvalue of A . For any matrix $A = [a_1, a_2, \dots, a_n] = (a_{ij}) \in \mathbb{R}^{m \times n}$ and $B = (b_{ij}) \in \mathbb{R}^{p \times q}$, the Kronecker product $A \otimes B$ is defined as $A \otimes B = (a_{ij}B) \in \mathbb{R}^{mp \times nq}$. We define $\text{vec}(A) = [a_1^T, a_2^T, \dots, a_n^T]^T \in \mathbb{R}^{mn}$. For a vector a , $\text{diag}(a)$ is a diagonal matrix whose diagonals are given as components of a . The remaining sections of this paper are organized as follows. Section 2 introduces some basic results. In section 3, we present our main perturbation results and show the mathematical equivalence of three kinds of condition numbers. We turn to the randomized algorithms in section 4 and the detailed error analysis for Algorithm RTTLS is given. The numerical results are performed in section 5 and section 6 concludes this paper.

2 Preliminaries

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ with $m \geq n$. Let $[A, b]$ and A have singular value decompositions, respectively

$$U^T[A, b]V = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_t) = \Sigma,$$

$$\widetilde{U}^T A \widetilde{V} = \text{diag}(\widetilde{\sigma}_1, \widetilde{\sigma}_2, \dots, \widetilde{\sigma}_n),$$

where $t = \min\{m, n+1\}$ and for the case $m > n$, orthonormal matrices U and V , diagonal matrix Σ are partitioned as follows:

$$U = [U_1, u_{n+1}]_{m \times (n+1)}, \quad V = \begin{bmatrix} V_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}_{(n+1) \times (n+1)}, \quad \Sigma = \begin{bmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \sigma_{n+1} \end{bmatrix}.$$

For the usual well-conditioned cases in this paper, we assume the genericity condition:

$$\widetilde{\sigma}_n > \sigma_{n+1}, \tag{2.1}$$

to ensure the existence and uniqueness of the TLS solution x (see [11]). The singular value σ_{n+1} can be treated as 0 for the case $m = n$ since σ_{n+1} does not exist at all. From best rank-1 approximation [12] of

matrix $[A, b]$, we know that

$$\begin{aligned} [E, f] &= -U \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_{n+1} \end{bmatrix} V^T \\ &= -\sigma_{n+1} u_{n+1} \begin{bmatrix} v_{12}^T & v_{22} \end{bmatrix} \\ &= -\sigma_{n+1} u_{n+1} v_{n+1}^T, \end{aligned}$$

where $v_{n+1} = [v_{12}^T, v_{22}]^T$. Therefore, $x = -v_{12}/v_{22}$. It follows from [38, Theorem 2.7] that the solution x can also be expressed as a function of $[A, b]$, i.e.,

$$x = (A^T A - \sigma_{n+1}^2 I)^{-1} A^T b, \quad (2.2)$$

and it holds that

$$\begin{bmatrix} x^T & -1 \end{bmatrix}^T = -\frac{1}{v_{22}} v_{n+1}. \quad (2.3)$$

In the case $m = n$ under the genericity condition, the original system is just a nonsingular one and the Tls solution is equal to the least squares solution. Now consider the Tls problem (1.1) and assume $\tilde{\sigma}_q > \sigma_{q+1} = \dots = \sigma_{n+1}$ with $q \leq n$. Let the above SVD still hold but partition V differently as follows:

$$V = \begin{bmatrix} V_{11} & V_{12} \\ v_{21} & v_{22} \end{bmatrix} \begin{matrix} n \\ 1 \end{matrix}. \quad (2.4)$$

$q \quad n+1-q$

The condition $\tilde{\sigma}_q > \sigma_{q+1} = \dots = \sigma_{n+1}$ is equivalent to that $\sigma_q > \sigma_{q+1} = \dots = \sigma_{n+1}$ and v_{22} is of full row rank, i.e., v_{22} is not a zero vector. According to [38, Theorem 3.10], the minimum norm Tls solution \bar{x} is given by

$$\bar{x} = -V_{12} v_{22}^\dagger = (V_{11}^T)^\dagger v_{21}^T.$$

This is called the truncated total least squares (Ttls). The case $m < n + 1$ requires that $\sigma_{n+1} = 0$ and hence $\|[E, f]\|_F = 0$ [38]. The idea of Ttls is to treat the small singular values of the augmented matrix $[A, b]$ as zeros and convert a numerically rank-deficient problem to an exactly rank-deficient one. For the discrete ill-posed problems where the singular values of the coefficient matrices decay gradually, Ttls can be applied, where the parameter q then plays the role of the regularization parameter. In practical applications, the smallest singular values of $[A, b]$ rarely coincide [39]. But if one considers the Tls problem as an approximation to the corresponding unobservable exact relation $A_0 x = b_0$, then $\text{rank}([A_0, b_0]) = \text{rank}(A_0) = q \leq n$. So $\sigma_{q+1}, \dots, \sigma_{n+1}$ are just the perturbations of zero. In this case it is realistic to define an error bound ϵ such that all singular values σ_i , satisfying $|\sigma_i - \sigma_{n+1}| < \epsilon$, are considered to coincide with σ_{n+1} . Therefore, we can use the formula $\bar{x} = -V_{12} v_{22}^\dagger$.

3 Perturbation results

First, we give a lemma which will be very useful in our analysis.

Lemma 3.1 Consider the total least squares problem (1.1) and assume that the genericity condition (2.1) holds. If $[A, b]$ is perturbed to $[A + \delta A, b + \delta b]$, then we have

$$\sigma_{n+1} u_{n+1}^T [\delta A, \delta b] v_{n+1} = \frac{r^T [\delta b - (\delta A)x]}{1 + x^T x},$$

where $r = b - Ax$.

PROOF. From (2.3) and the singular value decomposition of $[A, b]$, we know that

$$r = b - Ax = -[A, b] \begin{bmatrix} x \\ -1 \end{bmatrix} = \frac{1}{v_{22}} [A, b] v_{n+1} = \frac{1}{v_{22}} \sigma_{n+1} u_{n+1}.$$

Therefore we have

$$\begin{aligned} \frac{r^T [\delta b - (\delta A)x]}{1 + x^T x} &= \frac{\sigma_{n+1}}{v_{22}} \frac{u_{n+1}^T [\delta b - (\delta A)x]}{1 + x^T x} \\ &= -\sigma_{n+1} v_{22} u_{n+1}^T [\delta A, \delta b] \begin{bmatrix} x \\ -1 \end{bmatrix} \\ &= \sigma_{n+1} u_{n+1}^T [\delta A, \delta b] v_{n+1}, \end{aligned}$$

where we use $v_{22}^2 = \frac{1}{1+x^T x}$, which is a direct result of (2.3). □

The following lemma [37] is also needed for deriving our perturbation result.

Lemma 3.2 Let σ_{\min} be the smallest nonzero and simple singular value of a matrix X with u_{\min} and v_{\min} being its corresponding left and right singular vectors, respectively. If $\|\delta X\|_F$ is sufficiently small, then the smallest nonzero singular value $\widehat{\sigma}_{\min}$ of the perturbed matrix $\widehat{X} = X + \delta X$ is simple and

$$\widehat{\sigma}_{\min} = \sigma_{\min} + u_{\min}^T (\delta X) v_{\min} + O(\|\delta X\|_F^2).$$

In the following, we present our perturbation bound under the genericity condition (2.1).

Theorem 3.1 Consider the total least squares problem (1.1) and assume that the genericity condition (2.1) holds. If $\|[\delta A, \delta b]\|_F$ is sufficiently small, then we have that

$$\begin{aligned} \frac{\|\delta x\|_2}{\|x\|_2} &\lesssim \left(\frac{\|b\|_2}{\|x\|_2} \left\| (A^T A - \sigma_{n+1}^2 I)^{-1} A^T \right\|_2 \right) \frac{\|\delta b\|_2}{\|b\|_2} \\ &+ \left[\frac{\|A\|_2 \|r\|_2}{\|x\|_2} \left\| (A^T A - \sigma_{n+1}^2 I)^{-1} \right\|_2 + \|A\|_2 \left\| (A^T A - \sigma_{n+1}^2 I)^{-1} A^T \right\|_2 \right] \frac{\|\delta A\|_2}{\|A\|_2}, \end{aligned} \quad (3.1)$$

where $r = b - Ax$.

PROOF. From $(A^T A - \sigma_{n+1}^2 I)x = A^T b$, perturbing $[A, b]$ yields

$$\left[(A + \delta A)^T (A + \delta A) - \widehat{\sigma}_{n+1}^2 I_n \right] (x + \delta x) = (A + \delta A)^T (b + \delta b), \quad (3.2)$$

where $\widehat{\sigma}_{n+1}$ is the smallest singular value of $[A + \delta A, b + \delta b]$. Subtracting two equations (2.2) and (3.2), we have

$$\begin{aligned} (A^T A - \sigma_{n+1}^2 I_n) \delta x &= (\delta A)^T r + A^T [\delta b - (\delta A)x] + (\widehat{\sigma}_{n+1}^2 - \sigma_{n+1}^2)(x + \delta x) \\ &+ O(\|[\delta A, \delta b]\|_F^2). \end{aligned} \quad (3.3)$$

Furthermore, combining Lemma 3.1 and Lemma 3.2, we have

$$\begin{aligned} \widehat{\sigma}_{n+1}^2 - \sigma_{n+1}^2 &= (\widehat{\sigma}_{n+1} - \sigma_{n+1})(\widehat{\sigma}_{n+1} + \sigma_{n+1}) \\ &= \left\{ u_{n+1}^T [\delta A, \delta b] v_{n+1} + O(\|[\delta A, \delta b]\|_F^2) \right\} \left\{ 2\sigma_{n+1} + u_{n+1}^T [\delta A, \delta b] v_{n+1} + O(\|[\delta A, \delta b]\|_F^2) \right\} \\ &= 2 \frac{r^T [\delta b - (\delta A)x]}{1 + x^T x} + O(\|[\delta A, \delta b]\|_F^2), \end{aligned} \quad (3.4)$$

where for the last approximation we use Lemma 3.1. From (3.3) and (3.4), ignoring higher order terms, we know that

$$\begin{aligned} \delta x &\approx (A^T A - \sigma_{n+1}^2 I_n)^{-1} \left\{ \left[A^T (\delta b - (\delta A)x) + (\delta A)^T r \right] + 2 \frac{r^T [\delta b - (\delta A)x]}{1 + x^T x} (x + \delta x) \right\} \\ &\approx (A^T A - \sigma_{n+1}^2 I_n)^{-1} \left[A^T (\delta b) - A^T (\delta A)x + (\delta A)^T r \right] + 2 (A^T A - \sigma_{n+1}^2 I_n)^{-1} \frac{r^T [\delta b - (\delta A)x]}{1 + x^T x} x \\ &= (A^T A - \sigma_{n+1}^2 I_n)^{-1} \left[A^T + 2 \frac{x r^T}{1 + x^T x} \right] \delta b + (A^T A - \sigma_{n+1}^2 I_n)^{-1} \left[(\delta A)^T r - \left(A^T + 2 \frac{x r^T}{1 + x^T x} \right) (\delta A)x \right]. \end{aligned}$$

Hence,

$$\begin{aligned} \|\delta x\|_2 &\lesssim \left\| (A^T A - \sigma_{n+1}^2 I_n)^{-1} \left[A^T + 2 \frac{x r^T}{1 + x^T x} \right] \right\|_2 \|\delta b\|_2 \\ &+ \left\| (A^T A - \sigma_{n+1}^2 I_n)^{-1} \right\|_2 \|\delta A\|_2 \|r\|_2 + \left\| (A^T A - \sigma_{n+1}^2 I_n)^{-1} \left[A^T + 2 \frac{x r^T}{1 + x^T x} \right] \right\|_2 \|\delta A\|_2 \|x\|_2. \end{aligned}$$

Since $r = b - Ax = \frac{\sigma_{n+1}}{v_{22}} u_{n+1}$, using the MATLAB notation we have

$$\frac{x r^T}{1 + x^T x} = -v_{n+1}(1:n)\sigma_{n+1}u_{n+1}^T = -[V_{11}, v_{12}] \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_{n+1} \end{bmatrix} U^T.$$

Moreover, from the SVD of $[A, b]$, it follows that

$$A = [A, b] \begin{bmatrix} I_n \\ \mathbf{0}_{1 \times n} \end{bmatrix} = U \Sigma \begin{bmatrix} V_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}^T \begin{bmatrix} I_n \\ \mathbf{0}_{1 \times n} \end{bmatrix} = U \Sigma [V_{11}, v_{12}]^T, \quad \Sigma = \begin{bmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \sigma_{n+1} \end{bmatrix}.$$

Therefore, we obtain that

$$A^T + 2 \frac{x r^T}{1 + x^T x} = [V_{11}, v_{12}] \begin{bmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & -\sigma_{n+1} \end{bmatrix} U^T. \quad (3.5)$$

Furthermore,

$$\left[A^T + 2 \frac{x r^T}{1 + x^T x} \right] \cdot \left[A^T + 2 \frac{x r^T}{1 + x^T x} \right]^T = [V_{11}, v_{12}] \Sigma^2 [V_{11}, v_{12}]^T = A^T A,$$

and

$$\left\| \left(A^T A - \sigma_{n+1}^2 I_n \right)^{-1} \left[A^T + 2 \frac{x r^T}{1 + x^T x} \right] \right\|_2 = \left\| \left(A^T A - \sigma_{n+1}^2 I_n \right)^{-1} A^T \right\|_2. \quad (3.6)$$

Finally, we have

$$\begin{aligned} \frac{\|\delta x\|_2}{\|x\|_2} &\lesssim \left(\frac{\|b\|_2}{\|x\|_2} \left\| \left(A^T A - \sigma_{n+1}^2 I \right)^{-1} A^T \right\|_2 \right) \frac{\|\delta b\|_2}{\|b\|_2} \\ &+ \left[\frac{\|A\|_2 \|r\|_2}{\|x\|_2} \left\| \left(A^T A - \sigma_{n+1}^2 I \right)^{-1} \right\|_2 + \left\| \left(A^T A - \sigma_{n+1}^2 I \right)^{-1} A^T \right\|_2 \|A\|_2 \right] \frac{\|\delta A\|_2}{\|A\|_2}. \end{aligned}$$

□

The succinct perturbation bound above is based on the formula (3.6), which is derived by using (3.5) and the fact that $\|K\|_2^2 = \lambda_{\max}(KK^T)$ for any real matrix K . In fact, we can give another bound of the perturbation system, and express it as the following corollary.

Corollary 3.1 *Under the same conditions assumed in Theorem 3.1, we have*

$$\frac{\|\delta x\|_2}{\|x\|_2} \lesssim \left[\left\| \left(A^T A - \sigma_{n+1}^2 I_n \right)^{-1} A^T \right\|_2 \frac{\sqrt{1 + \|x\|_2^2}}{\|x\|_2} + \left\| \left(A^T A - \sigma_{n+1}^2 I_n \right)^{-1} \right\|_2 \frac{\|r\|_2}{\|x\|_2} \right] \|[\delta A, \delta b]\|_2. \quad (3.7)$$

PROOF. From the proof of Theorem 3.1, we know

$$\begin{aligned} \delta x &= \left(A^T A - \sigma_{n+1}^2 I_n \right)^{-1} \left(A^T + 2 \frac{x r^T}{1 + x^T x} \right) \delta b \\ &+ \left(A^T A - \sigma_{n+1}^2 I_n \right)^{-1} \left[(\delta A)^T r - \left(A^T + 2 \frac{x r^T}{1 + x^T x} \right) (\delta A) x \right] + \mathcal{O}(\|[\delta A, \delta b]\|_F^2), \end{aligned}$$

which can be rewritten as

$$\begin{aligned} \delta x &= \left(A^T A - \sigma_{n+1}^2 I_n \right)^{-1} \left[A^T + 2 \frac{x r^T}{1 + x^T x} \right] [\delta A, \delta b] \begin{bmatrix} -x \\ 1 \end{bmatrix} \\ &+ \left[\left(A^T A - \sigma_{n+1}^2 I_n \right)^{-1}, \mathbf{0}_{n \times 1} \right] [\delta A, \delta b]^T r + \mathcal{O}(\|[\delta A, \delta b]\|_F^2). \end{aligned}$$

Taking 2-norm on both sides, considering the property (3.6) and omitting the higher order terms, we simply get the bound for the relative error

$$\frac{\|\delta x\|_2}{\|x\|_2} \lesssim \left[\left\| \left(A^T A - \sigma_{n+1}^2 I_n \right)^{-1} A^T \right\|_2 \frac{\sqrt{1 + \|x\|_2^2}}{\|x\|_2} + \left\| \left(A^T A - \sigma_{n+1}^2 I_n \right)^{-1} \right\|_2 \frac{\|r\|_2}{\|x\|_2} \right] \|[\delta A, \delta b]\|_2.$$

□

Remark 1 In Theorem 3.1, if $m = n$, the relative error estimate (3.1) can be simplified as

$$\frac{\|\delta x\|_2}{\|x\|_2} \leq \frac{\|A^{-1}\|_2 \|b\|_2}{\|x\|_2} \frac{\|\delta b\|_2}{\|b\|_2} + \|A\|_2 \left(\|A^{-1}\|_2 + \|A^{-1}\|_2^2 \frac{\|r\|_2}{\|x\|_2} \right) \frac{\|\delta A\|_2}{\|A\|_2},$$

which is one specific case in the estimate of the least squares solution [32]. From the proof of Theorem 3.1, we know that

$$\begin{aligned} (A^T A - \sigma_{n+1}^2 I_n)^{-1} (\delta A)^T r &= (A^T A - \sigma_{n+1}^2 I_n)^{-1} A^T (A^T)^\dagger (\delta A)^T r \\ &= (A^T A - \sigma_{n+1}^2 I_n)^{-1} A^T \left[r^T \otimes (A^\dagger)^T \right] \text{vec}((\delta A)^T) \end{aligned}$$

and therefore the term $\left\| (A^T A - \sigma_{n+1}^2 I_n)^{-1} (\delta A)^T r \right\|_2 \leq \left\| (A^T A - \sigma_{n+1}^2 I_n)^{-1} A^T \right\|_2 \|r\|_2 \|A^\dagger\|_2 \|\delta A\|_F$. So we can get another bound

$$\|\delta x\|_2 \lesssim \left\| (A^T A - \sigma_{n+1}^2 I_n)^{-1} A^T \right\|_2 \left[\|\delta b\|_2 + \|A^\dagger\|_2 \|r\|_2 \|\delta A\|_F + \|x\|_2 \|\delta A\|_2 \right].$$

This bound is succinct but it is bigger than the bound in (3.1). It is easy to check that

$$\left\| (A^T A - \sigma_{n+1}^2 I)^{-1} A^T \right\|_2 = \frac{\tilde{\sigma}_n}{\tilde{\sigma}_n^2 - \sigma_{n+1}^2}.$$

We notice that the term $(A^T A - \sigma_{n+1}^2 I)^{-1} A^T$ also appears in the derivation of the “effective condition number” of the total least squares problem. The effective condition number is defined as [25, 26, 27]

$$\text{Cond}_{\text{eff}} = \frac{\|b\|_2}{\sigma_r \|x\|_2} = \frac{\|A^\dagger\|_2 \|b\|_2}{\|x\|_2}$$

for the linear system $Ax = b$ with σ_r being the smallest positive singular value of A . In some cases, the effective condition number is much smaller than the traditional one.

Denote

$$\begin{aligned} M &= \left[K \otimes b^T - x^T \otimes (KA^T) - K \otimes (Ax)^T, \quad KA^T \right], \\ N &= 2\sigma_{n+1} y \left(v_{n+1}^T \otimes u_{n+1}^T \right) \end{aligned}$$

with $K = (A^T A - \sigma_{n+1}^2 I)^{-1}$ and $y = Kx$. Omitting the complicated higher order term $R(\delta A, \delta b)$ in [44, Eqn.(3.5)], then the upper bound derived in [44] becomes

$$\frac{\|M + N\|_2 \| [A, b] \|_F}{\|x\|_2} \frac{\| [\delta A, \delta b] \|_F}{\| [A, b] \|_F},$$

and $K_{\text{ZLWQ}} = \frac{\|M+N\|_2 \| [A, b] \|_F}{\|x\|_2}$ can be defined as the condition number.

Denote $\tilde{D} = \text{diag}\left(\left(\tilde{\sigma}_1^2 - \sigma_{n+1}^2\right)^{-1}, \dots, \left(\tilde{\sigma}_n^2 - \sigma_{n+1}^2\right)^{-1}\right)$ and $D = \text{diag}\left(\sqrt{\sigma_1^2 + \sigma_{n+1}^2}, \dots, \sqrt{\sigma_n^2 + \sigma_{n+1}^2}\right)$. The upper bound obtained from [2] is expressed by

$$\sqrt{1 + \|x\|_2^2} \left\| \tilde{D} \begin{bmatrix} \tilde{V}^T, & \mathbf{0}_{n \times 1} \end{bmatrix} V \begin{bmatrix} D, & \mathbf{0}_{n \times 1} \end{bmatrix}^T \right\|_2 \frac{\|[A, b]\|_F}{\|x\|_2} \frac{\|[\delta A, \delta b]\|_F}{\|[A, b]\|_F},$$

and they define

$$K_{\text{BG}} = \sqrt{1 + \|x\|_2^2} \left\| \tilde{D} \begin{bmatrix} \tilde{V}^T, & \mathbf{0}_{n \times 1} \end{bmatrix} V \begin{bmatrix} D, & \mathbf{0}_{n \times 1} \end{bmatrix}^T \right\|_2 \frac{\|[A, b]\|_F}{\|x\|_2}$$

as the relative condition number.

Later, Li and Jia [24] established the following bound for the relative perturbation

$$K_{\text{L}} \frac{\|[\delta A, \delta b]\|_F}{\|[A, b]\|_F},$$

where

$$K_{\text{L}} = \frac{\left\| \left(A^T A - \sigma_{n+1}^2 I_n \right)^{-1} \left(2A^T \frac{r}{\|r\|_2} \frac{r^T}{\|r\|_2} G(x) - A^T G(x) + \begin{bmatrix} I_n \otimes r^T, & \mathbf{0}_{n \times m} \end{bmatrix} \right) \right\|_2 \|[A, b]\|_F}{\|x\|_2}$$

is the condition number with $G(x) = \begin{bmatrix} x^T, & -1 \end{bmatrix} \otimes I_m$.

We need to point out that, to derive the expressions for K_{LWQ} , K_{BG} and K_{L} , the higher order terms have been omitted in [44, 2, 24]. And it is reasonable to compare our bound given in Theorem 3.1 with the above three bounds. The numerical results will be given later.

Remark 2 Note that K_{BG} has another closed formula [2]

$$K_{\text{BG}} = \sqrt{1 + \|x\|_2^2} \left\| \left(A^T A - \sigma_{n+1}^2 I \right)^{-1} \left[A^T A + \sigma_{n+1}^2 \left(I_n - \frac{2xx^T}{1 + \|x\|_2^2} \right) \right] \left(A^T A - \sigma_{n+1}^2 I \right)^{-1} \right\|_2^{1/2} \frac{\|[A, b]\|_F}{\|x\|_2}.$$

Since $A^T A + \sigma_{n+1}^2 \left(I_n - \frac{2xx^T}{1 + \|x\|_2^2} \right) = A^T A - \sigma_{n+1}^2 I_n + 2\sigma_{n+1}^2 \left(I_n - \frac{xx^T}{1 + \|x\|_2^2} \right)$ is symmetric positive definite, we can define LL^T as its Cholesky factorization. Then we have

$$K_{\text{BG}} = \sqrt{1 + \|x\|_2^2} \left\| \left(A^T A - \sigma_{n+1}^2 I \right)^{-1} L \right\|_2 \frac{\|[A, b]\|_F}{\|x\|_2},$$

which is another expression of K_{L} [18].

Moreover, using Lemma 3.1 and the proof in [44, Lemma 3.2], we can get the following equation

$$M + N = \begin{bmatrix} -x^T \otimes D_{\sigma_{n+1}^2} + \left(A^T A - \sigma_{n+1}^2 I \right)^{-1} \otimes r^T, & D_{\sigma_{n+1}^2} \end{bmatrix},$$

where $D_{\sigma_{n+1}^2} = \left(A^T A - \sigma_{n+1}^2 I \right)^{-1} \left(A^T + 2 \frac{xr^T}{1 + x^T x} \right)$. Denote $P \in \mathbb{R}^{mn \times mn}$ the permutation matrix that represents the matrix transpose by $\text{vec}(B^T) = P \text{vec}(B)$. Note that K_{BG} can also be expressed by [2]

$$K_{\text{BG}} = \frac{\|\mathcal{M}_g\|_2 \|[A, b]\|_F}{\|x\|_2}$$

with

$$\mathcal{M}_{g'} = \begin{bmatrix} -x^T \otimes D_{\sigma_{n+1}^2} + \left(r^T \otimes (A^T A - \sigma_{n+1}^2 I)^{-1} \right) P, & D_{\sigma_{n+1}^2} \end{bmatrix}.$$

We can easily check that $M + N = \mathcal{M}_{g'}$, which means that $K_{BG} = K_{ZLWQ}$.

Therefore, we see that the condition numbers derived respectively in [2, 24, 44] are mathematically equivalent. But as pointed by the authors themselves, the normwise condition number proposed in [44] is not easy to compute.

4 Randomized algorithms for Tls problems

The randomized algorithms have been receiving increasingly more attention in numerical linear algebra, and they open the possibility of dealing with truly massive data sets, and have become more and more popular in the matrix approximation in the last decade [14]. Numerical experiments and detailed error analysis show that these random sampling techniques can be quite effective and more efficient than the classical competitors in many aspects. Avron et al. in [1] derived a randomized least-squares solver which outperforms LAPACK by large factors for dense highly overdetermined systems. Recently, Xiang and Zou [42] used the randomized strategy for solving large-scale discrete inverse problems. In this section, we first propose two algorithms for the cases where the numerical rank is known using the similar randomized strategies. One is the randomized algorithm for total least squares (RTLs for short), and the other is the randomized algorithm for truncated total least squares (RTLs for short). For the circumstances in which the target rank is not known, we further develop the adaptive randomized algorithms under the fixed precision (ARTLs for short). These randomized algorithms can greatly reduce the computational time, and still yield good approximate solutions.

4.1 Randomized algorithm RTLs for well-conditioned cases

(ALGORITHM RTLs: RANDOMIZED ALGORITHM FOR Tls)

1. Generate an $(n + 1) \times l$ Gaussian random matrix Ω .
 2. Solve $(C^T C) X = \Omega$, where $C = [A, b] \in \mathbb{R}^{m \times (n+1)}$.
 3. Compute the $(n + 1) \times l$ orthonormal matrix Q via QR factorization $X = QR$.
 4. Solve $(C^T C) Y = Q$.
 5. Form the $l \times l$ matrix $Z = Q^T Y = Q^T (C^T C)^{-1} Q$.
 6. Compute the SVD of the smaller symmetric matrix, $Z = W \Sigma W^T$, where W is orthogonal.
 7. Form the $(n + 1) \times l$ matrix $V = QW$, and define $v = V(:, 1)$.
 8. Form the solution $x_{RTLs} = -v(1 : n)/v(n + 1)$.
-

For the total least squares problem, it is very crucial to obtain the right singular vector v_{n+1} associated with the smallest singular values of $[A, b]$. Then the total least squares solution can be expressed by (2.3). For the expression (2.3), we know that the key point is to find the singular vector associated with the smallest singular value. But the randomized SVD [14] usually approximates well the largest singular values and the corresponding singular vectors. Suppose $C = [A, b]$ has the SVD $C = U\Sigma V^T$, where $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_{n+1})$, U and V are orthogonal matrices. If $\sigma_{n+1} = 0$, then b is in the range of A and the TLS solution is equal to the least squares solution. We do not consider this trivial case here. Then $C^T C = V\Sigma^T \Sigma V^T$, and $(C^T C)^{-1} = V \text{diag}(\sigma_{n+1}^{-2}, \dots, \sigma_1^{-2}) V^T$. Hence we can see that σ_{n+1}^{-2} becomes the largest diagonal element, and we can apply the randomized algorithm to approximate this value and achieve its corresponding singular vector v_{n+1} . The essential step of this traditional algorithm is the SVD of $[A, b]$. But when the size of A is large, SVD can be very costly, or even prohibitive. How to reduce the computational cost and still ensure the accuracy of the approximate solution is our main concern. Our new randomized algorithm RTLS is presented in Algorithm RTLS.

Note that l is a pre-specified parameter. In [14] the index l is usually selected in the form $l = k + p$, where p is an oversampling parameter, and k corresponds to the rank k specified in advance for the best rank- k approximation of A . To understand Algorithm RTLS more, we make some remarks about each step of the algorithm. In Step 2 we obtain $X = (C^T C)^{-1} \Omega$ to extract the column information, which is further represented by an orthogonal matrix Q in Step 3. The linear system involving $C^T C$ in Step 2 and 4 can be solved by direct methods such as Cholesky factorization or Krylov subspace iterative methods. When the problem is not too ill-conditioned, this coefficient matrix is symmetric positive definite, and can be solved quite efficiently. After Step 5 the problem is reduced to a smaller symmetric semi-positive definite matrix $Z = Q^T (C^T C)^{-1} Q$, and SVD is applied to this small matrix in Step 6. This leads to an SVD approximation, $(C^T C)^{-1} \approx V\Sigma V^T$, where $V = QW$ and $W\Sigma W^T = Z$. We then use this approximate SVD to seek the approximate total least squares solution x_{RTLS} in Step 8.

4.2 Randomized algorithm RTLS for ill-conditioned cases

Algorithm RTLS works well for the well-conditioned cases. For the total least squares problem with very ill-conditioned coefficient matrices, the condition number of $C^T C$ can be very large since the condition number $\text{Cond}(C^T C) = \text{Cond}(C)^2$. We need to use regularization techniques to avoid noise contaminations and obtain a meaningful approximate solution. Fierro et al. in [9] focused on the truncated TLS for solving discrete ill-posed problems, where the singular values of the coefficient matrix decay gradually. The technique of truncated TLS is similar in spirit to truncated SVD (TsVD), where the small singular values of $[A, b]$ are treated as zeros, and the problem is reduced to an exactly rank-deficient one [9]. Recently, the sensitivity analysis and conditioning has been given in [13] and some applications of the truncated TLS are reported [13]: System identification, linear system theory, image reconstruction, speech and audio processing, modal and spectral analysis, chemometrics, computer vision, machine learning, computer algebra, and astronomy. The traditional truncated total least squares solution is given by the following Algorithm TLS [38, Section 3.6.1].

In Algorithm TLS, the truncation parameter k is user-specified or determined adaptively [9]. It is chosen such that the first k large singular values dominate and $\|v_{22}\|_2 \neq 0$. Here the Moore-Penrose inverse $v_{22}^\dagger = v_{22}^T \|v_{22}\|_2^{-2}$.

(ALGORITHM TTLS: CLASSICAL TRUNCATED TLS)

1. Compute the SVD: $[A, b] = U\Sigma V^T = \sum_{i=1}^{n+1} \sigma_i u_i v_i^T$, where $A \in \mathbb{R}^{m \times n}$.
 2. Partition the matrix, $V = \begin{bmatrix} V_{11} & V_{12} \\ v_{21} & v_{22} \end{bmatrix}$, where $V_{12} \in \mathbb{R}^{n \times (n+1-k)}$, $v_{21} \in \mathbb{R}^{1 \times k}$, and $v_{22} \in \mathbb{R}^{1 \times (n+1-k)}$.
 3. Form the minimum-norm TLS solution: $x_{\text{TTLS}} = -V_{12} v_{22}^\dagger$.
-

When the discrete ill-posed problems is of medium size, we can compute the complete SVD of $[A, b]$ directly like Step 1 in Algorithm TTLS. When the size of A is large, the SVD in Step 1 is very costly since the SVD needs about $6mn^2 + 20n^3$ flops [12]. This flaw leads us to improve the efficiency by computing the SVD of $[A, b]$ in Step 1 “partially.” The corresponding algorithm is named “partial total least squares (PTLS)” in [38]. The only difference between TTLS and PTLS lies in the Step 1: one uses the classical complete SVD, while the other one applies the partial SVD. The authors in [38] report that PTLS is two times faster than TTLS while the same accuracy can be maintained. Moreover, the relative efficiency of partial SVD increases when the dimension of the desired singular subspace is relatively smaller to the dimension n . For large-scale discrete ill-posed problems, Lanczos bi-diagonalization in [9] is used to achieve a good approximation to the singular triplets associated with several largest singular values. This approach will lose the sparsity or structure of the coefficient matrix in the first step of bi-diagonal reduction. What’s more, Lanczos procedure needs to access the coefficient matrix many times and use the BLAS-2 operations, i.e., the matrix-vector multiplications. Here we propose an alternative technique based on randomized strategies, that is, a randomized version of truncated total least squares (RTTLS). This is a new randomized algorithm, most flops spent on the matrix-matrix multiplications, which are the so-called nice BLAS-3 operations, and the algorithm can be realized by accessing the original large-scale matrix A only once.

(ALGORITHM RTTLS: RANDOMIZED ALGORITHM FOR TRUNCATED TLS)

1. Generate an $(n+1) \times l$ Gaussian random matrix Ω .
 2. Form the $m \times l$ matrix $Y = C\Omega$, where $C = [A, b]$.
 3. Apply QR decomposition to Y , i.e., $Y = QR$, where $Q \in \mathbb{R}^{m \times l}$.
 4. Form the $l \times (n+1)$ matrix Z such that $Z = Q^T C$.
 5. Apply SVD to the smaller matrix Z , i.e., $Z = W\Sigma V^T$, where $V \in \mathbb{R}^{(n+1) \times l}$.
 6. Let $V_{11} = V(1:n, 1:k)$, $v_{21} = V(\text{end}, 1:k)$, and form the solution $x_{\text{RTTLS}} = (V_{11}^T)^\dagger v_{21}^T$.
-

Usually the randomized algorithm cannot approximate the small singular values very well, hence we do not prefer to use the expression $x_{\text{TTLS}} = -V_{12} v_{22}^\dagger$ directly. Since in Algorithm RTTLS we can

obtain a good approximation of the right singular vectors associated with largest singular values, we use $x_{\text{RTTLS}} = (V_{11}^T)^\dagger v_{21}^T$ in Step 6. In Algorithm RTTLS the parameter l stands for the number of sampling, and the number k is the parameter for truncating ($k \leq l$). A larger l will improve the reliability of the algorithm [14], but also increase the computational complexity. In practice, we choose $l \ll n$, and make a balance between the reliability and the computational complexity. The truncation parameter k can be user-specified or determined by some regularization technique if no a priori estimate. Here we use randomized regularization techniques in [42] to obtain an estimation for this parameter. We first perform randomized algorithms to obtain an approximate SVD of A , then a Gcv function based on this approximation is used to determine the truncation parameter k for the TsVD solution of $Ax \approx b$. This procedure can be performed very fast [42]. This parameter cannot be the optimal for the total least squares based on the SVD of the augmented matrix $[A, b]$, but should be a reasonable estimate for the truncation parameter in TTLS. Other rules such as the L-curve, quasi-optimality, and discrepancy principle can be also used for regularization parameter choice. Our randomized TTLS is constructed in the spirit of the truncated SVD (TsVD). Tikhonov regularization for TTLS [4, 10, 21, 23, 29] can be also combined with randomized algorithms, together with the existing rules for regularization parameter choice, such as L-curves, Gcv, quasi-optimality, and discrepancy principle, etc. The detailed discussion about some important issues, such as the regularization parameter choice, the scaling of A and b [38, Section 3.6.2], is beyond the scope of this paper.

Step	RTLS	RTTLS	TTLS	PTLS
1	$O(nl)$	$O(nl)$	$6mn^2 + 20n^3$	$O(mnl)$
2	$2mn^2 + \frac{2}{3}n^3$	$2mnl$	-	-
3	$4nl^2 - \frac{4}{3}l^3$	$4ml^2 - \frac{4}{3}l^3$	$O(n^2 - nk)$	$O(n^2 - nk)$
4	$\frac{2}{3}n^3$	$2mnl$	-	-
5	$2nl^2$	$6nl^2 + 20l^3$	-	-
6	$26l^3$	$O(nk^2)$	-	-
7	$2nl^2$	-	-	-
8	$O(n)$	-	-	-

Table 1: Computational complexity.

4.3 Adaptive randomized algorithms for truncated TTLS

The randomized algorithms discussed above are used to solve the fixed-rank problems. In practical applications, the target rank is rarely known in advance. We do not need to determine it accurately. So the adaptive approach [14] is usually implemented to increase the number of samples until the error $\|C - QQ^T C\|_2$ satisfies the desired tolerance. The tolerance parameter is a standard for measuring whether the basis matrix Q captures the action of the target matrix C . The theoretical basis behind this scheme is that we can estimate the exact error $\|C - QQ^T C\|_2$ by computing $\|(I - QQ^T)C\omega\|_2$ with ω

being a standard Gaussian vector. Draw r standard Gaussian vectors, then

$$\|(I - QQ^T)C\|_2 \leq 10\sqrt{\frac{2}{\pi}} \max_{1 \leq i \leq r} \|(I - QQ^T)C\omega_i\|_2$$

holds except with probability $1 - 10^{-r}$ where r is an integer that balances computational cost and reliability [14].

Given the augmented matrix $C = [A, b]$, a tolerance ϵ , and integer r , the formal schemes for computing an orthonormal basis in Step 3 of Algorithm RTLS and therefore finding the truncated solution are described in Algorithm ARTLS. It follows that $\|C - QQ^TC\|_2 \leq \epsilon$ holds with probability at least $1 - \min\{m, n+1\}10^{-r}$. We need to stress that the reorthogonalization is implemented in Step 6 and Step 7 to overcome the numerical instability that the column vectors of Q become small as increasing the basis. The CPU time requirements of Algorithms ARTLS and RTLS are essentially identical [14].

4.4 Computational complexity

We shall say a few words about the computational complexity of the randomized algorithms. The cost of each step of algorithms is listed in Table 1. In Table 1, we denote the PTLs the corresponding Algorithm TILS using the Lanczos bi-diagonalization to fulfill the partial SVD in Step 1. As discussed in the above subsection, we have to form the solution by exploiting the singular vectors corresponding to the largest singular values, i.e., $x_{\text{PTLS}} = (V_{11}^T)^\dagger v_{21}^T$. For the matrix $C = [A, b] \in \mathbb{R}^{m \times (n+1)}$, the flops count of the classical SVD based on R-bidiagonalization is about $6mn^2 + 20n^3$ [12], while the cost of Algorithm RTLS is about $2mn^2 + \frac{4}{3}n^3 + 8nl^2 + O(l^3)$; the cost of Algorithm RTLS is about $4mnl + (4m + 6n)l^2 + O(l^3)$. The cost of Algorithm RTLS is much cheaper than the classical one. Note that the most flops are performed in Step 2 of RTLS by very efficient BLAS-3 operations, and that fast Krylov subspace iterative solvers can be used in Step 2 and 4 of RTLS instead. We can see that the computational cost of PTLs is of the same magnitude as RTLS. But PTLs just carries out BLAS-2 operations and it will be not as efficient as it looks in practical computations. The advantage of our RTLS can be more obvious than just what the flops account tells.

For the cases where singular values decay rapidly, we can choose a small parameter l . For most cases, $m \gtrsim n \gg l$. According to the flops, the ratio of the cost for Algorithm RTLS over that for the classical SVD is of the order $O(l/n)$. Hence, the randomized algorithms can be essentially faster than the traditional counterpart.

4.5 Error estimates

We will analyze the accuracy of the Algorithm RTLS in this part. Before the main results, we introduce an important estimate in [14].

Lemma 4.1 [14, Corollary 10.9] *Suppose that $A \in \mathbb{R}^{m \times n}$ has singular values $\sigma_1 \geq \sigma_2 \geq \dots$. Choose a target rank $k \geq 2$ and an oversampling parameter $p \geq 4$, where $k + p \leq \min\{m, n\}$. Draw an $n \times (k + p)$ standard Gaussian matrix Ω and let Q be an orthonormal matrix whose columns form a basis for the range of the sampled matrix $A\Omega$. Then*

$$\|A - QQ^TA\|_2 \leq (1 + 9\sqrt{k+p}\sqrt{\min\{m, n\}})\sigma_{k+1},$$

(ALGORITHM ARTTLS: ADAPTIVE RANDOMIZED ALGORITHM FOR TRUNCATED TLS)

1. Generate standard Gaussian random vectors $\omega_1, \dots, \omega_r$ of length n .
 2. For $i = 1, \dots, r$, compute $y_i = C\omega_i$.
 3. Set $j = 0$ and $Q^{(0)} = [\]$, i.e., the $m \times 0$ empty matrix.
 4. while $\max \{\|y_{j+1}\|_2, \dots, \|y_{j+r}\|_2\} \geq \epsilon / (10 \sqrt{2/\pi})$,
 5. $j = j + 1$.
 6. Overwrite y_j by $\left[I - Q^{(j-1)} (Q^{(j-1)})^T \right] y_j$.
 7. $q_j = y_j / \|y_j\|_2$.
 8. $Q^{(j)} = \left[Q^{(j-1)}, q_j \right]$.
 9. Draw a standard Gaussian random vector ω_{j+r} of length n .
 10. $y_{j+r} = \left[I - Q^{(j)} (Q^{(j)})^T \right] C\omega_{j+r}$.
 11. $\left[y_{j+1}, \dots, y_{j+r-1} \right] = \left[y_{j+1}, \dots, y_{j+r-1} \right] - q_j q_j^T \left[y_{j+1}, \dots, y_{j+r-1} \right]$.
 12. end while
 13. $Q = Q^{(j)}$.
 14. Form the $j \times (n+1)$ matrix $Z = Q^T C$.
 15. Apply SVD to the smaller matrix Z , i.e., $Z = W\Sigma V^T$, where $V \in \mathbb{R}^{(n+1) \times j}$.
 16. Let $V_{11} = V(1:n, 1:j)$, $v_{21} = V(\text{end}, 1:j)$, and form the solution $x_{\text{ARTTLS}} = \left(V_{11}^T \right)^\dagger v_{21}^T$.
-

with failure probability at most $3p^{-p}$.

From the process of Algorithm `RTTLS`, we see that $USV^T = QW\Sigma V^T = QQ^TC$ where we denote $U = QW$ and $C = [A, b]$. So $\|C - USV^T\|_2 = \|C - QQ^TC\|_2$. Hence we obtain a good SVD approximation for C with high probability.

Before studying the accuracy of the stochastic procedures in the algorithm, we review the perturbation results given by Wei [39, Theorem 4.1], which is stated in the following slightly modified lemma.

Lemma 4.2 *Consider the TLS problem (1.1). Let the SVD for A and $[A, b]$ be given as in the preliminaries. Assume that for some $q \leq n$, $\tilde{\sigma}_q > \sigma_{q+1}$. Partition V as in (2.4), let $\widehat{A} \in \mathbb{R}^{m \times n}$, $\widehat{b} \in \mathbb{R}^m$, and $[\widehat{A}, \widehat{b}] = [A, b] + E$ with $\|E\|_2 \leq \frac{1}{6}(\tilde{\sigma}_q - \sigma_{q+1})$, and the SVD for $[\widehat{A}, \widehat{b}]$ be*

$$\widehat{U}^T [\widehat{A}, \widehat{b}] \widehat{V}^T = \widehat{\Sigma}.$$

Partition \widehat{V} conformally with V and replace V_{ij} by \widehat{V}_{ij} for $i, j = 1, 2$. Define $\widehat{x}_{\text{TLS}} = (\widehat{V}_{11}^T)^\dagger \widehat{V}_{21}^T$ and $x_{\text{TLS}} = (V_{11}^T)^\dagger V_{21}^T$. When $x_{\text{TLS}} \neq \mathbf{0}$, the following estimate holds:

$$\frac{\|x_{\text{TLS}} - \widehat{x}_{\text{TLS}}\|_2}{\|x_{\text{TLS}}\|_2} \leq \frac{12(\|E\|_2 + \sigma_{q+1})}{\tilde{\sigma}_q - \sigma_{q+1}} \frac{\sigma_1}{\|b\|_2 - \sigma_{q+1}}.$$

Using Lemma 4.1 and Lemma 4.2, we get the estimate below.

Theorem 4.1 *Assume $m \geq n + 1$. Assume that $[A, b]$ has singular values $\sigma_1, \dots, \sigma_{n+1}$ and A has singular values $\tilde{\sigma}_1 \geq \dots \geq \tilde{\sigma}_n$. Moreover, assume $\tilde{\sigma}_q > \sigma_{q+1}$ with $q \leq n$ and let k be the target rank of $[A, b]$ and let x_{RTTLS} be the approximate TLS solution by performing Algorithm `RTTLS` with the Gaussian random matrix $\Omega \in \mathbb{R}^{n \times (k+p)}$. If*

$$\sigma_{k+1} \leq \frac{\tilde{\sigma}_q - \sigma_{q+1}}{6 + 54\sqrt{(k+p)n}},$$

then we have

$$\frac{\|x_{\text{TLS}} - x_{\text{RTTLS}}\|_2}{\|x_{\text{TLS}}\|_2} \leq \frac{12\sigma_1 \left[(1 + 9\sqrt{(k+p)n})\sigma_{k+1} + \sigma_{q+1} \right]}{(\tilde{\sigma}_q - \sigma_{q+1})(\|b\|_2 - \sigma_{q+1})} \quad (4.1)$$

with failure probability at most $3p^{-p}$. More specifically, if $q = k$ is the numerical rank of $[A, b]$, we get the bound below with probability not less than $1 - 3p^{-p}$

$$\frac{\|x_{\text{TLS}} - x_{\text{RTTLS}}\|_2}{\|x_{\text{TLS}}\|_2} \leq \frac{12\sigma_1 (2 + 9\sqrt{(k+p)n})}{\tilde{\sigma}_k \|b\|_2} \sigma_{k+1} + \mathcal{O}(\sigma_{k+1}^2). \quad (4.2)$$

PROOF. Denote that $C = [A, b]$ and $\widehat{C} = QQ^TC$. From Lemma 4.1 and the assumption we know that

$$\|C - \widehat{C}\|_2 = \|C - QQ^TC\|_2 \leq (1 + 9\sqrt{(k+p)n})\sigma_{k+1} \leq \frac{1}{6}(\tilde{\sigma}_q - \sigma_{q+1})$$

with probability not less than $1 - 3p^{-p}$. Then applying Lemma 4.2 we obtain (4.1). For the specific cases, if the numerical rank of $[A, b]$ is $k = q$, it means that σ_{k+1} is very close to zero. The bound in (4.1) can be simplified to

$$\frac{\|x_{\text{TTLs}} - x_{\text{RTTLs}}\|_2}{\|x_{\text{TTLs}}\|_2} \leq \frac{12\sigma_1 (2 + 9\sqrt{(k+p)n})}{(\widetilde{\sigma}_k - \sigma_{k+1})(\|b\|_2 - \sigma_{k+1})} \sigma_{k+1}.$$

Consider the Taylor expansion for the function $f(x) = 1/[(\widetilde{\sigma}_k - x)(\|b\|_2 - x)]$ at $x = 0$, we obtain that

$$f(\sigma_{k+1}) = \frac{1}{\widetilde{\sigma}_k \|b\|_2} + \left(\frac{1}{\widetilde{\sigma}_k^2 \|b\|_2} + \frac{1}{\widetilde{\sigma}_k \|b\|_2^2} \right) \sigma_{k+1} + \mathcal{O}(\sigma_{k+1}^2).$$

Substituting this equation into the above inequality directly, we can get (4.2). \square

We point out that the assumption for σ_{k+1} usually holds for the ill-conditioned cases, where $k = q$ is the numerical rank and we treat the other smaller singular values as zeros. During these cases, the upper bound (4.2) is of order $\mathcal{O}(\sigma_{k+1})$ and hence the relative error of the solution from RTTLs and the solution from TTLs is small.

5 Numerical examples

In this section we give numerical examples to verify the perturbation bounds and our randomized total least squares algorithms (RTLs and RTTLs). The following numerical tests are performed via MATLAB R2010a in a laptop with Intel Core i5 by using double precision.

5.1 Perturbation bounds

We compare our upper bounds (3.1) and (3.7) with those derived in [2, 24, 44]. We will see that these three are equal, and ours are sharper.

Example I. In this example [2, Example 1] we consider the TLs problem $Ax \approx b$, where $[A, b]$ is defined by

$$[A, b] = Y \begin{bmatrix} D \\ \mathbf{0} \end{bmatrix} Z^T \in \mathbb{R}^{m \times (n+1)}, Y = I_m - 2yy^T, Z = I_{n+1} - 2zz^T,$$

where $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^{n+1}$ are random unit vectors, $D = \text{diag}(n, n-1, \dots, 1, 1 - \epsilon_p)$ for a given parameter ϵ_p . The quantity $\widetilde{\sigma}_n - \sigma_{n+1}$ measures the distance of our problem to nongenericity and, due to the interlacing property, we have in exact arithmetic

$$\widetilde{\sigma}_n - \sigma_{n+1} \leq \sigma_n - \sigma_{n+1} = \epsilon_p.$$

We consider a random perturbation $\|[\delta A, \delta b]\|_F = 10^{-10}$. We take $m = 100$, $n = 40$ in this example and denote $\Delta = \frac{\|[\delta A, \delta b]\|_F}{\|[A, b]\|_F}$.

In Table 2, we compare the exact relative error with the upper bounds (3.1) and the above bounds derived in [2, 24, 44]. Without considering the computational cost, we can see that the numerical results of the three condition numbers in [2, 24, 44] are the same. We observe that our bounds are sharp and smaller than the bounds derived in the literature.

ϵ_p	$\frac{\ \tilde{x}-x\ _2}{\ x\ _2}$	$K_{\text{ZLWQ}}\Delta$	$K_{\text{BG}}\Delta$	$K_{\text{LJ}}\Delta$	(3.1)	(3.7)
9.99976032E-1	2.6233E-11	1.5815E-09	1.5815E-09	1.5815E-09	2.8145E-10	3.9565E-10
9.99952397E-5	3.8714E-07	1.1343E-05	1.1343E-05	1.1343E-05	3.2472E-06	3.8752E-06

Table 2: Comparisons of forward error and upper bounds for a perturbed TLS problem.

m	$\frac{\ \tilde{x}-x\ _2}{\ x\ _2}$	$K_{\text{ZLWQ}}\Delta$	$K_{\text{BG}}\Delta$	$K_{\text{LJ}}\Delta$	(3.1)	(3.7)
100	1.1553E-13	1.0152E-11	1.0152E-11	1.0152E-11	4.7548E-12	4.6281E-12
250	2.5302E-14	6.3627E-12	6.3627E-12	6.3627E-12	1.9277E-12	1.9001E-12

Table 3: Comparisons of forward error and upper bounds for a perturbed TLS problem.

Example II. Consider the second example from [38, p. 42], where

$$A = \begin{bmatrix} m-1 & -1 & \cdots & -1 \\ -1 & m-1 & \cdots & -1 \\ \vdots & & & \\ -1 & -1 & \cdots & m-1 \\ -1 & -1 & \cdots & -1 \\ -1 & -1 & \cdots & -1 \end{bmatrix} \in \mathbb{R}^{m \times (m-2)}, \quad b = \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \\ m-1 \\ -1 \end{bmatrix} \in \mathbb{R}^m.$$

The exact solution of the TLS problem $Ax \approx b$ is $x = -[1, 1, \dots, 1]^T$ and $\sigma_{n+1} = \sqrt{m}$, $\tilde{\sigma}_n = \sqrt{2m}$. We consider the same random perturbations as the Example I. The results are listed in Table 3. From the experience of computing, we also find that the bound in [44] is quite impractical for computing, since MATLAB will be out of memory on our Microsoft Windows operating system.

5.2 Numerical experiments for randomized algorithms

In this subsection, we apply Algorithm R_{TLS}, Algorithm RT_{TLS} and Algorithm AR_{TLS} to Example I, Example II and some cases in Hansen's Regularization Tool [15]. We will compare the computational time and solution accuracy of our new randomized TLS algorithms with the traditional algorithms.

5.2.1 Algorithm R_{TLS} on well-conditioned cases

For the case Example I in Table 4, we choose $\epsilon_p = 9.99976031\text{e-}1$, and set $n = \frac{2}{5}m$. The solution x_{TLS} is computed by (2.3), while x_{RTLS} is obtained by Algorithm R_{TLS}. Denote the relative error $\text{Err}_{\text{RTLS}} = \|x_{\text{TLS}} - x_{\text{RTLS}}\|_\infty / \|x_{\text{TLS}}\|_\infty$. The corresponding execution time Time_{TLS} and $\text{Time}_{\text{RTLS}}$ are measured by the MATLAB tic-toc pairs in seconds. From Table 4, we can see that our R_{TLS} algorithm on large matrices outperforms the traditional counterpart according to computational time, while the accuracy of solutions of two methods is comparable. For the small matrices, the advantage of R_{TLS} will not be so obvious. For an ill-conditioned matrix, MATLAB reports inaccuracy warning due to the ill-conditioned linear system in Step 2 and 4 of Algorithm R_{TLS}. Even for the ill-conditioned case DERIV2, Algorithm R_{TLS} can still give approximate solution with good accuracy. But for the very ill-conditioned cases, we need Algorithm RT_{TLS}.

	Matrix size	Cond(A)	Cond([A, b])	Time _{TLS}	Time _{RTLS}	Err _{RTLS}
Example I	$m = 500$	2.00E+2	8.34E+6	0.0698	0.0110	6.48E-10
	$m = 1000$	4.00E+2	1.67E+7	0.5643	0.0941	1.06E-10
	$m = 5000$	2.00E+3	8.34E+7	35.021	2.7903	2.40E-09
Example II	$m = 500$	15.8	22.4	0.2063	0.0402	5.53E-02
	$m = 1000$	22.4	31.6	1.3648	0.2706	4.09E-02
	$m = 5000$	50.0	70.7	154.51	23.345	1.88E-02
Deriv2	$m = 500$	3.04E+5	3.33E+5	0.3180	0.0477	5.34E-05
	$m = 1000$	1.22E+6	1.33E+6	1.7050	0.4033	6.56E-04
	$m = 5000$	3.04E+7	3.33E+7	157.64	39.114	5.15E-01

Table 4: Tests on RTLS ($l = 10$). Time_{TLS} and Time_{RTLS} are the computational times (in seconds) for the algorithms TLS and RTLS respectively. The relative error Err_{RTLS} = $\|x_{\text{TLS}} - x_{\text{RTLS}}\|_{\infty} / \|x_{\text{TLS}}\|_{\infty}$.

λ_j	c_j
$-0.082 \pm 0.926i$	1
$-0.147 \pm 2.874i$	1
$-0.188 \pm 4.835i$	1
$-0.220 \pm 6.800i$	1
$-0.247 \pm 8.767i$	1
$-0.270 \pm 10.733i$	1

Table 5: Six pairs of poles and residues.

5.2.2 Examples based on TLS-Prony modeling

The TLS approach is a promising method in the field of signal processing. Rahman and Yu [34] presented a method for frequency estimation using TLS for solving the linear prediction equation. The problem here is taken from [30]. We first consider a set of linear prediction equations. Assume $a_j = [y_{j-1}, \dots, y_{j+m-2}]^T$ where $y_l = \sum_{j=1}^p c_j z_j^l$, $z_j = \exp(\lambda_j T)$, $j = 1, \dots, p$. The λ_j 's and c_j 's are to be determined. Furthermore, assume c_j and z_j are nonzeros and z_j 's are distinct for $j = 1, \dots, p$. Let $A_n = [a_1, \dots, a_n]$, $b_n = -a_{n+1}$ and consider the linear system

$$A_n x = b_n. \quad (5.1)$$

Assume $m \geq n$, $m \geq p$. It is known [41] that $\text{rank}(A_n) = \min\{n, p\}$. So if $n \geq p$, then (5.1) is compatible. For any solution $x = (\alpha_0, \alpha_1, \dots, \alpha_{n-1})^T$, construct a polynomial

$$P_n(z) = z^n + \alpha_{n-1}z^{n-1} + \dots + \alpha_1 z + \alpha_0,$$

then we know that P_n has zeros z_1, \dots, z_p . We choose λ_j and c_j as in Table 5. In this example $T = 0.2$, $m = 2000$, $p = 12$, $n = 1000$ are used and we compare the TLS with the RTLS where the sampling size is chosen as $l = p + 1$. The plots for the solutions are shown in Figure 1 and the infinity norm relative error is $6.7623e - 8$, while the time for TLS using partial SVD and RTLS are 0.8924 seconds and 0.0333 seconds respectively.

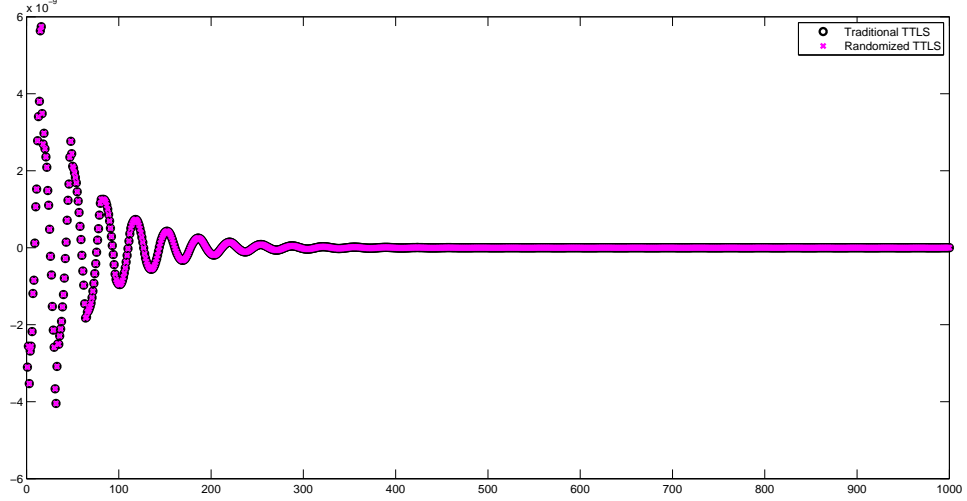


Figure 1: Computed solutions for the true TLS solution and the randomized one.

δ	k	Time _{TTLS}	Time _{PTLS}	Time _{RTTLS}	Err _{RTTLS}
1E-1	3	0.0123	0.2230	0.0039	8.04E-3
1E-2	5	0.0133	0.2365	0.0039	8.92E-4
1E-3	7	0.0119	0.2455	0.0075	1.59E-3
1E-4	8	0.0114	0.2424	0.0039	3.76E-4

Table 6: Tests on SHAW with different relative noise levels. The relative error $\text{Err}_{\text{RTTLS}} = \|x_{\text{RTTLS}} - x_{\text{TTLS}}\|_{\infty} / \|x_{\text{TTLS}}\|_{\infty}$. Algorithm RTTLS is substantially faster than TTLS and PTLS.

5.2.3 Algorithm RTTLS on ill-conditioned cases

Our ill-conditioned cases are taken from Hansen's Regularization Tools [15]. For example, the case SHAW is generated by the command $[\bar{A}, \bar{b}, x_{\text{true}}] = \text{SHAW}(m)$. Then noises are added to \bar{A} and \bar{b} . Suppose that δ is the relative noise level. We define

$$b = \bar{b} + \delta \|\bar{b}\|_2 \frac{\zeta}{\|\zeta\|_2}, \quad A = \bar{A} + \delta \|\bar{A}\|_F \frac{Z}{\|Z\|_F},$$

where ζ is a random vector, $\zeta = 2 * \text{rand}(m, 1) - 1$; Z is a random matrix, $Z = 2 * \text{rand}(m) - 1$. It is easy to verify that

$$\frac{\|b - \bar{b}\|_2}{\|\bar{b}\|_2} = \frac{\|A - \bar{A}\|_F}{\|\bar{A}\|_F} = \delta.$$

Then we seek the total least squares solution of $Ax \approx b$.

We first test Algorithm RTTLS on the 100×100 matrix SHAW with different relative noise levels δ . The results are given in Table 6. The truncation parameter k is estimated by the randomized algorithm with

Gcv and Tsvd [42]. After the determination of parameter k , the computational time for implementing Algorithm TTLS and Algorithm RTTLS is recorded in $\text{Time}_{\text{TTLS}}$ and $\text{Time}_{\text{RTTLS}}$ respectively. And $\text{Time}_{\text{PTLS}}$ denotes the time cost in TTLS using Lanczos bi-diagonalization based partial Svd. Here we denote the relative error $\text{Err}_{\text{RTTLS}} = \|x_{\text{TTLS}} - x_{\text{RTTLS}}\|_{\infty} / \|x_{\text{TTLS}}\|_{\infty}$. From our computing, we see that the computed solutions of Algorithm PTLS are almost the same as those of Algorithm TTLS, and hence the relative errors for the solutions of PTLS which we denote as $\text{Err}_{\text{PTLS}} = \|x_{\text{PTLS}} - x_{\text{TTLS}}\|_{\infty} / \|x_{\text{TTLS}}\|_{\infty}$ are much smaller. Here we ignore the error Err_{PTLS} and do not list it in the table. From this table we can see that the results of Algorithm RTTLS are very close to those of traditional TTLS even for the relative noise level as large as 10%. According to the computational time, Algorithm PTLS does not show obvious advantages over TTLS for small size cases, while Algorithm RTTLS is substantially faster than the traditional TTLS. The computed solutions for the case where the relative noise level $\delta=1\text{E-}3$ are presented in Figure 2.

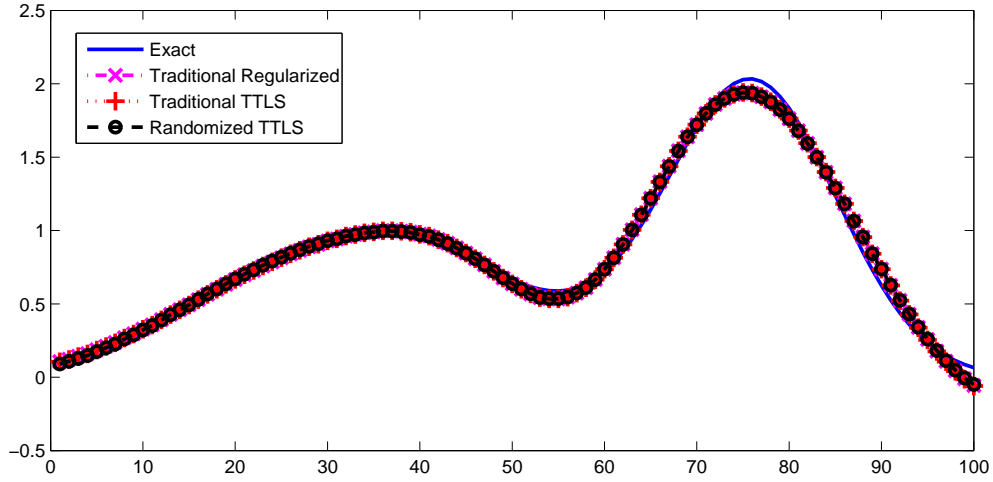


Figure 2: Computed solutions for the case SHAW of size $m=100$ with relative noise level $\delta=1\text{E-}3$.

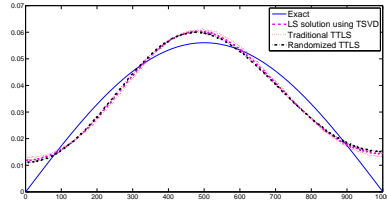
We then test Algorithm RTTLS on larger matrices. We set the parameter for sampling size $l = 10$ and the relative noise level $\delta=1\text{E-}3$ for all cases. The results are given in Table 7. The marker * in Table 7 means we cannot load the example LLaplace on our computer when the size $n = 5000$. Obviously, PTLS can be much faster than the TTLS when the size of the matrix becomes larger. But it is still not as efficient as the randomized one because of its BLAS-2 operations. The randomized strategy can greatly speed up the classical Algorithm TTLS. The advantage of our Algorithm RTTLS is more obvious when we test the larger matrices. The plots of the computed solutions are given in Figure 3.

5.2.4 Test on Adaptive Algorithm ARTTLS

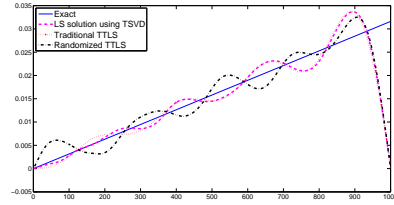
We still use the examples from Hansen's Regularization Tools [15] and test the case with matrix size $m = n = 1000$. Here we set $r = 7$ in the Algorithm ARTTLS. Different tolerances generate different j 's in Algorithm ARTTLS. So we tried several ϵ 's to make sure that k 's in Algorithm RTTLS and j 's in Algorithm ARTTLS are close, and then the comparisons for the relative errors and time are reasonable.

	Matrix size	k	Time _{TLS}	Time _{PTLS}	Time _{RTLS}	Err _{RTLS}
Baart	$m = 100$	4	0.0153	0.2907	0.0040	6.43E-3
	$m = 1000$	4	1.7561	0.2664	0.0143	6.53E-3
	$m = 5000$	4	176.47	1.5042	0.2471	5.86E-3
Deriv2	$m = 100$	6	0.0130	0.2604	0.0040	1.39E-2
	$m = 1000$	7	1.6727	0.3589	0.0148	6.96E-2
	$m = 5000$	9	170.40	1.8398	0.2506	1.20E-2
Foxgood	$m = 100$	2	0.0129	0.2691	0.0037	4.60E-6
	$m = 1000$	3	1.7638	0.2795	0.0143	5.09E-4
	$m = 5000$	3	171.88	1.1383	0.2227	1.14E-4
Gravity	$m = 100$	7	0.0152	0.4679	0.0039	1.91E-3
	$m = 1000$	8	1.7214	0.2963	0.0147	6.70E-3
	$m = 5000$	9	183.92	2.2156	0.3014	3.16E-2
Heat	$m = 100$	8	0.0107	0.2283	0.0041	7.33E-2
	$m = 1000$	9	1.7172	0.3963	0.0163	3.93E-2
	$m = 5000$	9	165.56	1.4494	0.2551	8.15E-2
LLaplace	$m = 100$	8	0.0182	0.2972	0.0056	2.22E-4
	$m = 1000$	9	3.4499	0.5609	0.0410	1.83E-2
	$m = 5000$	*	*	*	*	*
Phillips	$m = 100$	7	0.0109	0.2476	0.0038	1.66E-3
	$m = 1000$	7	2.3627	0.2804	0.0137	2.24E-3
	$m = 5000$	7	174.74	1.1844	0.2194	6.08E-3

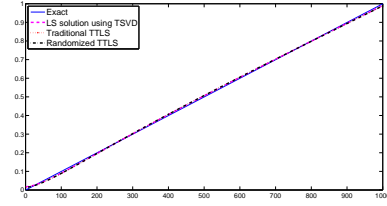
Table 7: Algorithm RTLS on ill-conditioned cases.



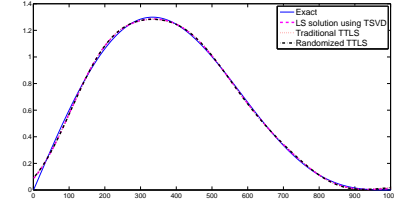
(a) BAART



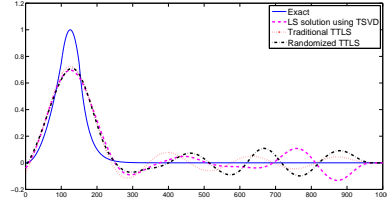
(b) DERIV2



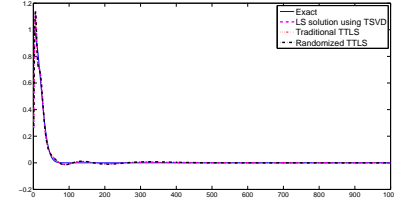
(c) FOXGOOD



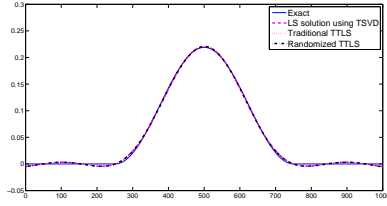
(d) GRAVITY



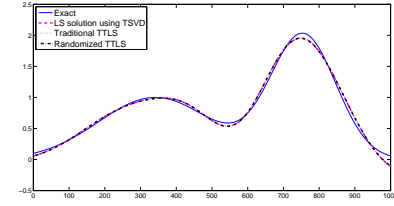
(e) HEAT



(f) ILAPLACE



(g) PHILLIPS



(h) SHAW

Figure 3: RTTLS for ill-conditioned cases of size $m = 1000$ with relative noise level $\delta=1E-3$.

	k	ϵ	j	Time _{P_{TLS}}	Time _{R_{TLS}}	Time _{ARTLS}	Err _{R_{TLS}}	Err _{ARTLS}
Baart	4	8E-1	4	0.4970	0.0266	0.0567	2.22E-2	4.24E-2
Deriv2	7	2E-2	8	0.4062	0.0223	0.0335	4.39E-2	1.45E-1
Foxgood	3	5E-1	3	0.3086	0.0206	0.0302	3.48E-4	2.14E-3
Gravity	8	7E-1	8	0.3449	0.0246	0.0436	5.66E-3	9.82E-3
Heat	9	4E-1	8	0.3412	0.0275	0.0383	2.08E-1	7.03E-2
LLaplace	9	7E-1	10	0.6054	0.0408	0.1226	5.38E-3	7.07E-2
Phillips	7	4E-0	8	0.3218	0.0213	0.0367	1.44E-3	7.80E-3
Shaw	7	6E-1	7	0.6624	0.0330	0.0487	3.40E-3	1.03E-2

Table 8: Algorithm ARTLS on ill-conditioned cases. The relative errors $\text{Err}_{\text{RTLS}} = \|x_{\text{RTLS}} - x_{\text{TLS}}\|_{\infty} / \|x_{\text{TLS}}\|_{\infty}$, $\text{Err}_{\text{ARTLS}} = \|x_{\text{TLS}} - x_{\text{ARTLS}}\|_{\infty} / \|x_{\text{TLS}}\|_{\infty}$. Both RTLS and ARTLS need less computational time than PTLS based on Lanczos procedure.

The performance for Algorithm ARTLS is shown in Table 8. In the table, $\text{Err}_{\text{ARTLS}}$ denotes the relative error $\|x_{\text{TLS}} - x_{\text{ARTLS}}\|_{\infty} / \|x_{\text{TLS}}\|_{\infty}$ and $\text{Time}_{\text{ARTLS}}$ represents the time cost for Algorithm ARTLS. It is clear that Algorithm ARTLS can still give good accuracy with less computational time than the traditional one under the fixed precision.

6 Conclusion

In this paper, we derive a new perturbation bound for the total least squares problem. This sharper and numerically computable perturbation bound is well illustrated by the numerical examples. Also we show that three kinds of condition numbers in [2, 24, 44] obtained through different ways are mathematically equivalent. We propose randomized algorithms RTLS, RTLS and ARTLS for the numerical solutions of well-conditioned and ill-conditioned total least squares problems, respectively. These randomized algorithms can greatly reduce the computational time, and still give solutions with good accuracy. The regularization parameter in RTLS is estimated by the truncated parameter of the TsVD solution of $Ax \approx b$ based on a fast randomized SVD of A [42]. Then a randomized SVD of $[A, b]$ together with this truncation parameter yields a good approximate TLS solutions to the large-scale ill-conditioned total least squares problems. The detailed investigation on other regularization parameter choices, and other techniques such as Tikhonov regularization, will be our future research.

7 Acknowledgments

The authors would like to thank Marc Baboulin, Ken Hayami, Zhongxiao Jia, Lothar Reichel, and Jun Zou for their comments and suggestions which led to improvements of our manuscript.

References

- [1] Blendenpik: Supercharging LAPACK’s least-squares solver, Avron, Haim and Maymounkov, Petar and Toledo, Sivan, SIAM J. Sci. Comput., Vol. 32, No. 3, pp. 1217–1236.

- [2] Baboulin, Marc and Gratton, Serge, A contribution to the conditioning of the total least-squares problem, *SIAM J. Matrix Anal. Appl.*, Vol. 32, 2011, No. 3, pp. 685–699.
- [3] Baboulin, Marc and Gratton, Serge and Lacroix, Rémi and Laub, Alan, Efficient computation of condition estimates for linear least squares problems. To appear in the Proceedings of the 10th International Conference on Parallel Processing and Applied Mathematics, PPAM 2013(09/2013).
- [4] Beck, Amir and Ben-Tal, Aharon, On the solution of the Tikhonov regularization of the total least squares problem, *SIAM J. Optim.*, Vol. 17, 2006, No. 1, pp. 98–118.
- [5] Beck, Amir and Ben-Tal, Aharon and Teboulle, Marc, Finding a global optimal solution for a quadratically constrained fractional quadratic problem with applications to the regularized total least squares, *SIAM J. Matrix Anal. Appl.*, Vol. 28, 2006, No. 2, pp. 425–445.
- [6] Chang, X.-W. and Tittley-Peloquin, D., Backward perturbation analysis for scaled total least-squares problems, *Numer. Linear Algebra Appl.*, Vol. 16, 2009, No. 8, pp. 627–648.
- [7] De Moor, Bart and David, Johan, Total linear least squares and the algebraic Riccati equation, *Systems Control Lett.*, Vol. 18, 1992, No. 5, pp. 329–337.
- [8] Diao, Huaian and Shi, Xinghua and Wei, Yimin, Effective condition numbers and small sample statistical condition estimation for the generalized Sylvester equation, *Sci. China Math.*, Vol. 56, 2013, No. 5, pp. 967–982.
- [9] Fierro, R. D. and Golub, G. H. and Hansen, P. C. and O’Leary, D. P., Regularization by truncated total least squares, *SIAM J. Sci. Comput.*, Vol. 18, 1997, No. 4, pp. 1223–1241.
- [10] Golub, Gene H. and Hansen, Per Christian and O’Leary, Dianne P., Tikhonov regularization and total least squares, *SIAM J. Matrix Anal. Appl.*, Vol. 21, 1999, No. 1, pp. 185–194 .
- [11] Golub, Gene H. and Van Loan, Charles F., An analysis of the total least squares problem, *SIAM J. Numer. Anal.*, Vol. 17, 1980, No. 6, pp. 883–893.
- [12] Golub, Gene H. and Van Loan, Charles F., *Matrix Computations*, Johns Hopkins Studies in the Mathematical Sciences, 4th, Johns Hopkins University Press, Baltimore, MD, 2013, pp. xxx+698.
- [13] Gratton, Serge and Tittley-Peloquin, David and Ilunga, Jean Tshimanga, Sensitivity and conditioning of the truncated total least squares solution, *SIAM J. Matrix Anal. Appl.*, Vol. 34, 2013, No. 3, pp. 1257–1276.
- [14] Halko, N. and Martinsson, P. G. and Tropp, J. A., Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions, *SIAM Rev.*, Vol. 53, 2011, No. 2, pp. 217–288.
- [15] Hansen, Per Christian, Regularization tools: a Matlab package for analysis and solution of discrete ill-posed problems, *Numer. Algorithms*, Vol. 6, 1994, No. 1-2, pp. 1–35.
- [16] Hansen, Per Christian and Nagy, James G and O’leary, Dianne P, *Deblurring Images: Matrices, Spectra, and Filtering*, Vol. 3, 2006, Society for Industrial and Applied Mathematics, Philadelphia.
- [17] Hnětynková, Iveta and Plešinger, Martin and Sima, Diana Maria and Strakoš, Zdeněk and Van Huffel, Sabine, The total least squares problem in $AX \approx B$: a new classification with the relationship to the classical works, *SIAM J. Matrix Anal. Appl.*, Vol. 32, 2011, No. 3, pp. 748–770.
- [18] Jia, Zhongxiao and Li, Bingyu, On the condition number of the total least squares problem, *Numer. Math.*, Vol. 125, 2013, No. 1, pp. 61–87.
- [19] Kenney, C. S. and Laub, A. J., Small-sample statistical condition estimates for general matrix functions, *SIAM J. Sci. Comput.*, Vol. 15, 1994, No. 1, pp. 36–61.

- [20] Kenney, C. S. and Laub, A. J. and Reese, M. S., Statistical condition estimation for linear least squares, *SIAM J. Matrix Anal. Appl.*, Vol. 19, 1998, No. 4, pp. 906–923.
- [21] Lampe, J. and Voss, H., Global convergence of RTLSQEP: a solver of regularized total least squares problems via quadratic eigenproblems, *Math. Model. Anal.*, Vol. 13, 2008, No. 1, pp. 55–66.
- [22] Lampe, Jörg and Voss, Heinrich, Large-scale Tikhonov regularization of total least squares, *J. Comput. Appl. Math.*, Vol. 238, 2013, pp. 95–108.
- [23] Lee, Geunseop and Fu, Haoying and Barlow, Jesse L, Fast High-Resolution Image Reconstruction Using Tikhonov Regularization Based Total Least Squares, *SIAM J. Sci. Comput.*, Vol. 35, No. 1, pp. B275–B290, 2013.
- [24] Li, Bingyu and Jia, Zhongxiao, Some results on condition numbers of the scaled total least squares problem, *Linear Algebra Appl.*, Vol. 435, 2011, No. 3, pp. 674–686.
- [25] Li, Zi-Cai and Chien, Cheng-Sheng and Huang, Hung-Tsai, Effective condition number for finite difference method, *J. Comput. Appl. Math.*, Vol. 198, 2007, No. 1, pp. 208–235.
- [26] Li, Zi-Cai and Huang, Hung-Tsai and Chen, Jeng-Tzong and Wei, Yimin, Effective condition number and its applications, *Computing*, Vol. 89, 2010, No. 1-2, pp. 87–112.
- [27] Li, Zi-Cai and Huang, Hung-Tsai and Wei, Yimin, Ill-conditioning of the truncated singular value decomposition, Tikhonov regularization and their applications to numerical partial differential equations, *Numer. Linear Algebra Appl.*, Vol. 18, 2011, No. 2, pp. 205–221.
- [28] Liu, Xin-Guo, Solvability and perturbation analysis of the total least squares problem, *Acta Math. Appl. Sinica*, Vol. 19, 1996, No. 2, pp. 254–262.
- [29] Lu, Shuai and Pereverzev, Sergei V. and Tautenhahn, Ulrich, Regularized total least squares: computational aspects and error bounds, *SIAM J. Matrix Anal. Appl.*, Vol. 31, 2009, No. 3, pp. 918–941.
- [30] Majda, George and Strauss, Walter A. and Wei, Musheng, Computation of exponentials in transient data, *IEEE Transactions on Antennas and Propagation*, 37 (1989), pp. 1284C1290.
- [31] Markovsky, Ivan and Rastello, Maria Luisa and Premoli, Amedeo and Kukush, Alexander and Van Huffel, Sabine, The element-wise weighted total least-squares problem, *Comput. Statist. Data Anal.*, Vol. 50, 2006, No. 1, pp. 181–209.
- [32] Malyshev, A. N., A unified theory of conditioning for linear least squares and Tikhonov regularization solutions, *SIAM J. Matrix Anal. Appl.*, Vol. 24, 2003, No. 4, pp. 1186–1196.
- [33] Paige, Christopher C. and Strakoš, Zdeněk, Scaled total least squares fundamentals, *Numer. Math.*, Vol. 91, 2002, No. 1, pp. 117–146.
- [34] Rahman, MD and Yu, Kai-Bor, Total least squares approach for frequency estimation using linear prediction, *IEEE Transactions on Acoustics, Speech and Signal Processing*, 35 (1987), pp. 1440–1454.
- [35] Renaut, Rosemary A. and Guo, Hongbin, Efficient algorithms for solution of regularized total least squares, *SIAM J. Matrix Anal. Appl.*, Vol. 26, 2004, No. 2, pp. 457–476.
- [36] Sima, Diana M. and Van Huffel, Sabine and Golub, Gene H., Regularized total least squares based on quadratic eigenvalue problem solvers, *BIT*, Vol. 44, 2004, No. 4, pp. 793–812.
- [37] Stewart, G. W., A second order perturbation expansion for small singular values, *Linear Algebra Appl.*, Vol. 56, 1984, pp. 231–235.

- [38] Van Huffel, Sabine and Vandewalle, Joos, The Total Least Squares Problem, *Frontiers in Applied Mathematics*, Vol. 9, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1991, pp. xiv+300.
- [39] Wei, Musheng, The analysis for the total least squares problem with more than one solution, *SIAM J. Matrix Anal. Appl.*, Vol. 13, 1992, No. 3, pp. 746–763.
- [40] Wei, Musheng, Algebraic relations between the total least squares and least squares problems with more than one solution, *Numer. Math.*, Vol. 62, 1992, No. 1, pp. 123–148.
- [41] Wei, Musheng and Majda, George, A new theoretical approach for Prony’s method, *Linear Algebra Appl.*, 136(1990), pp.119-132.
- [42] Xiang, Hua and Zou, Jun, Regularization with randomized SVD for large-scale discrete inverse problems, *Inverse Problems*, Vol. 29, 2013, No. 8, pp. 085008, 23.
- [43] Xu, Wei and Qiao, Sanzheng and Wei, Yimin, A note on the scaled total least squares problem, *Linear Algebra Appl.*, Vol. 428, 2008, No. 2-3, pp. 469–478.
- [44] Zhou, Liangmin and Lin, Lijing and Wei, Yimin and Qiao, Sanzheng, Perturbation analysis and condition numbers of scaled total least squares problems, *Numer. Algorithms*, Vol. 51, 2009, No. 3, pp. 381–399.